between the cerebral hemispheres in recognizing faces presented from different viewpoints, and I came across a paper that had looked at the effects of pose on facial recognition.

Laboratory studies had suggested that a three-quarter view leads to better recognition, but Logie *et al.* wanted to know if this generalized to the real world. In their first experiment they replicated some of the laboratory findings. So far so good. Their second experiment was to publish in a local newspaper, photographs of six different faces from different viewpoints, instruct readers that the six people would be walking around Cambridge at a certain time the following Saturday, and invite the readers to identify them and contact the experimenters by phone or by filling in a response form from the newspaper. Only one of the 100 000 readers responded. The experimenters then published a request for readers to say whether or not they had participated in the task. No responses were received.

This was an important study for me. A well-respected group of researchers had carried out a well-planned, somewhat extravagant experiment and it had turned to dust in their hands. If it could happen to them, it could happen to anyone. Undeterred, they carried out a third experiment in which volunteers were paid to find the targets after viewing photographs. This time the weather intervened: "Unfortunately, on the night prior to the experiment Cambridge received an unusually heavy fall of snow." As a result, almost 25% of the paid volunteers didn't take part, and only two reported accurate sightings.

The experiment was failing for a good reason — the sheer difficulty of recognizing multiple targets from photographs — and the authors eventually solved this problem, but it is for the documentation of the failed experiments and the circumstances which led to failure that I remember this paper. It was the first lesson I received in how to be wrong, and whenever the first two stabs at an experiment don't work out, it's to this paper that I turn for reassurance.

References

1. Logie, RH, Baddeley, AD, Woodhead, MM: Face recognition, pose and ecological validity. *Appl Cogn Psychol* 1987, 1:53–69.

Address: Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK.

Correspondence

The *Drosophila* Beat protein is related to adhesion proteins that contain immunoglobulin domains

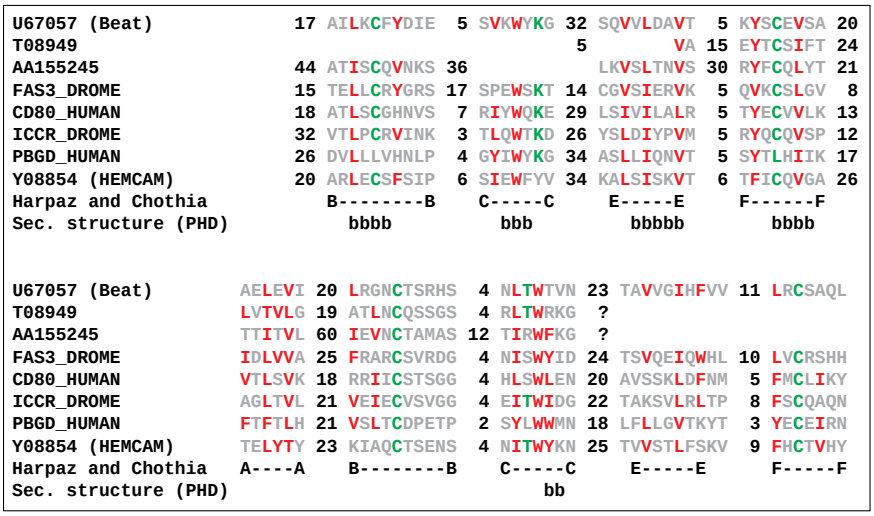
Arcady R. Mushegian

Axon extension along the surface of other axons in the course of embryo development results in axon bundle formation (fasciculation), followed at the later stages by controlled defasciculation of axon subsets at defined choice points. The genetic and biochemical basis of fasciculation in chick and fruitfly embryos has been studied in some detail, and cell adhesion molecules (CAMs) have been shown to have a pivotal role in this process. One well-studied example is Fasciclin II (Fas II) of *Drosophila*, a membrane CAM containing immunoglobulin-like domains that enables axon fasciculation along the three longitudinal axon pathways [1]. Post-translational modifications (by phosphorylation and sialylation) of Fas II and of related chick CAMs, have been implicated in the control of fasciculation, presumably by affecting the extent of homophilic adhesion [2,3].

Recently, a *Drosophila* gene, *beaten path*, has been shown to promote defasciculation by decreasing axon–axon adhesion. Motor axons fail to defasciculate in *beaten path* mutant embryos, and this phenotype is reversed in embryos in which both *beaten path* and genes encoding CAMs are mutant, consistent with the antagonistic roles of these two classes of proteins [4]. Beat, the product of the *beaten path* gene, is a secreted protein that was thought not to contain any known sequence motifs [4].

We applied the sensitive and selective BLAST2 algorithm [5] to search the non-redundant sequence database at NCBI (Bethesda, USA) with the Beat sequence. This search retrieved many classes of proteins with known immunoglobulin-like domains, including chicken CAMs, the T-lymphocyte activation protein CD80, the irregular chiasm C-roughest protein required for correct axonal pathway formation in the optic lobe of *Drosophila*, and *Drosophila* Fas III protein. The matches were in the amino-terminal half of Beat protein, and corresponded to the predicted immunoglobulin-like domains of the other proteins. Although probability of matching by chance was of moderate significance (p values of 10^{-2}), the typical pairwise scores between Beat and other proteins were in the range of 90–100, indicating likely biological relevance of the observed similarities for a medium-sized protein with unbiased sequence composition [6]. Moreover, searching of the expressed sequence tags (EST) database with Beat sequence detected the putative products of a human and a mouse EST — T08949 and AA155245 — that were even more closely related to Beat, and displayed statistically significant similarity to this protein (scores 118 and 109, and probabilities of matching by chance 10^{-5} and 10^{-4}, respectively). The EST database search also retrieved many immunoglobulin-domain proteins at statistically significant levels when used as queries in the further rounds

Figure 1



Immunoglobulin-like domains in the Beat protein. The ungapped blocks of high sequence similarity, for which the β-strand assignment [8] was possible, are shown. The independent, single-sequence prediction of the β strands in the Beat sequence with the PHD program [14] with accuracy of 7 or above is also shown. The unique identifier of each sequence in SWISSPROT or GenBank is given to the left of each sequence. Distances, in amino-acid residues, between blocks and to the ends of the sequences are indicated by numbers. Protein sequences encoded by human and mouse ESTs are conceptual translations made by the TBLASTN program, and the numbers stand for the distances in nucleotides. Conserved bulky hydrophobic residues (I, L, V, M, W, Y, and F) are shown in red, other conserved residues are shown in green. The alignment was constructed using the MACAW program [15].

of database searches. Multiple sequence alignment detected two immunoglobulin-like domains in the Beat sequence (Fig. 1), with invariant residues and predicted secondary structure conforming to the known pattern of sequence conservation in diverse immunoglobulin-like domains [7,8]. Some of the similarities between Beat and the immunoglobulin-domain proteins could also be observed using the conventional BLAST algorithm [9], after filtering out the carboxy-terminal, low-sequence-complexity segment of Beat protein with the SEG program [10]. Thus, Beat seems to be a secreted immunoglobulin-domain protein.

Interestingly, the putative immunoglobulin-like domains of Beat also matched regions in a group of human pregnancy zone proteins, not previously reported to be immunoglobulin-related. Database search using pregnancy zone protein sequences revealed that they, too, contain multiple immunoglobulin-like domains. For example, the probability of matching by chance between human pregnancy-specific β-1 glycoprotein D and human B-cell receptor CD22B, a *bona fide* immunoglobulin-domain CAM [11], was computed, by BLAST2 program, to be less then 10^{-9}, with a similarity score of 150. The complete sets of immunoglobulin domains in the pregnancy zone proteins and related carcinoembryonic antigens remains to be determined.

Surprisingly, the carboxy-terminal half of Beat protein immediately downstream of the immunoglobulin-like domains matched, in BLAST2 search, the human paraoxonase sequence (similarity score 97; data not shown). The catalytic mechanism and structure–function relationships of paraoxonase are unclear [12], and it is not yet possible to predict whether Beat possesses paraoxonase-like enzymatic activity. The presence of the tandem immunoglobulin-like domains in Beat defasciculation protein has implications for its function. It has been suggested that FasII and other immunoglobulin-domain CAMs promote fasciculation by homophilic and heterophilic adhesion, mediated by their immunoglobulin domains [1,13]. A secreted immunoglobulin domain might bind by homophilic adhesion to the membrane-bound immunoglobulin domain of CAM, thereby reducing its interaction with other cell surfaces. Detailed structural modelling and direct biochemical experimentation will answer the question of whether Beat indeed promotes defasciculation by CAM displacement.

References

1. Lin DM, Fetter RD, Kopczynski C, Grenningloh G, Goodman CS: Genetic analysis of Fasciclin II in ***Drosophila***: defasciculation, refasciculation, and altered fasciculation. *Neuron* 1994, 13:1055–1069.
2. Rutishauser U, Landmesser L: Polysialic acid on the surface of axons regulates patterns of normal and activity-dependent innervation. *Trends Neurosci* 1991, 14:528–532.
3. Desai CJ, Gindhardt JG Jr, Goldstein LS, Zinn K: Receptor tyrosine phosphatases are required for motor axon guidance in the ***Drosophila*** embryo. *Cell* 1996, 84:599–609.
4. Fambrough D, Goodman CS: The ***Drosophila beaten path*** gene encodes a novel secreted protein that regulates defasciculation at motor axon choice points. *Cell* 1996, 87:1049–1058.
5. Altschul SF, Gish W: Local alignment statistics. *Methods Enzymol* 1996, 266:460–480.
6. Altschul SF, Boguski MS, Gish W, Wootton JC: Issues in searching molecular sequence databases. *Nature Genet* 1994, 6:119–129.
7. Bork P, Holm L, Sander C: The immunoglobulin fold. Structural classification, sequence patterns, and common core. *J Mol Biol* 1994, 242:309–320.
8. Harpaz Y, Chothia C: Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. *J Mol Biol* 1994, 238:528–539.
9. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. *J Mol Biol* 1990, 215:403–410.
10. Wootton JC, Federhen S: Analysis of

compositionally biased regions in sequence databases. *Methods Enzymol* 1996, 266:554–573.
11. Wilson GL, Fox CH, Fauci AS, Kerhl JH: cDNA cloning of the B cell membrane protein CD22: a mediator of B–B cell interactions. *J Exp Med* 1991, 173:137–146.
12. Sorenson RC, Primo-Parmo SL, Kuo C-L, Adkins S, Lockridge O, La Du BN: Reconsideration of the catalytic center and mechanism of mammalian paraoxonase/arylesterase. *Proc Natl Acad Sci USA* 1995, 92:7187–7191.
13. Stoeckli ET, Landmesser LT: Axonin-1, Nr-CAM, and Ng-CAM play different roles in the ***in vivo*** guidance of chick comissural neurons. *Neuron* 1995, 14:1165–1179.
14. Rost B: PHD: predicting one-dimensional protein structure by profile-based neural networks. *Methods Enzymol* 1996, 266:525–539.
15. Schuler GD, Altschul SF, Lipman DJ: A workbench for multiple alignment construction and analysis. *Proteins Struct Funct Genet* 1991, 9:180–190.

Address: Sequana Therapeutic, Inc., 11099 North Torrey Pines Road, La Jolla, California 92037, USA.

Modular structure of the ***Drosophila*** Beat protein

J. Fernando Bazan* and Corey S. Goodman†

Neuronal growth cones are often guided by, and extend along, the surface of other axons, forming axon bundles or fascicles. This process of fasciculation can be highly discriminating, providing labeled axon pathways to steer growth cones towards particular regions [1]; in turn, growth cones selectively defasciculate, branching from the main axon highway at specific choice points. The genetic analysis of these processes in *Drosophila* has begun to shed light on the molecular mechanisms that control the selective entry and exit of growth cones from axon pathways, and to implicate key regulatory genes (see [2] for review).

Critical insights have emerged about the nature and polarity of guidance signals that migrating cells and axons receive from their extracellular environment during neural development. These cues can be contact-mediated, or take the form of diffusible chemical messengers. Moreover, they can be either attractive or repulsive. The balance of these directional 'forces' regulates a variety of guidance decisions, including fasciculation and defasciculation [2]. Accordingly, the surfaces of axons and other neural cells are studded with molecular sensors — proteins acting as adhesive antennae or sensitive receptors — that show a remarkable uniformity of architecture [3], regardless of whether they respond to contact-mediated or long-range signals [2].

A model system to study this molecular choreography arises in the shaping of the motor nerve pathways in the *Drosophila* embryo, where subsets of motor axons branch off at particular choice points and then steer into their muscle target regions [4]. This process of defasciculation appears to require the downmodulation of the function of Fasciclin II (Fas II) — an adhesive, immunoglobulin (Ig) superfamily counterpart of vertebrate NCAM [5] — without affecting its expression on axonal surfaces. Genetic analyses thus far have identified five genes as potent negative regulators of Fas II [4,6,7]. One of these genes — *beaten path* (*beat* [4]) — was recently shown to encode a novel secreted protein that is expressed by motor neurons [7]. The Beat protein probably regulates defasciculation by autocrine stimulation of an unidentified axonal receptor, as direct inhibition of Fas II is not observed [7].

We have used sensitive methods to reanalyze the sequence of Beat, and located a number of protein motifs that provide new insights into its structure, function and evolution. The mature 401 amino-acid chain of Beat can be reliably parsed into three discrete globular domains: a tandem repeat of Ig modules with distinctive Cys and Trp landmarks (D1, residues 1–114; D2, 115–220) separated by an unstructured linker (221–293) from a Cys-rich carboxy-terminal segment (D3; 294–401) (Fig. 1a). Database searches with WU-BLAST2 (W. Gish, unpublished; http://blast.wustl.edu) show that the amino-terminal Ig D1 module preferentially draws vertebrate T-cell receptor sequences, whereas the more economical Ig D2 elicits matches with more divergent Ig superfamily molecules like *Drosophila* Fasciclin III (Fig. 1a). Overlay of predicted secondary structures [8,9] and fold recognition analysis [10] propose a comparative, structure-based classification [11,12] of the Beat Ig domains as nine-β-stranded V-type structures; the D2 module displays shorter loops [12].

The carboxy-terminal D3 module contains a pattern of six Cys residues loosely similar to a 'cysteine-knot' motif conserved in the β-sheet folds of growth factors like transforming growth factor-β, neurotrophins (for example, nerve growth factor), platelet-derived growth factors and glycoprotein hormones [13]. In particular, the sequences **C**NMS**G**R**C** (residues 332–338) and **C**G**C** (385–387) in Beat match, respectively, the **C**X$_{1–9}$**G**X**C** and **C**X**C** cysteine-knot fingerprints (as detected by PRINTS motif GFCYSKNOT [14]), typically separated by ~40–60 amino acids (Fig. 1b). Three intra-chain disulfide links in a knot-like topology are predicted in Beat (Cys297–Cys370, Cys332–Cys385 and Cys338–Cys387); a seventh, unpaired Cys396 may be available for an intermolecular link. This latter suggestion is enhanced by the observed propensity of cysteine-knot domains to form homodimers or heterodimers [13]. This latent function of cysteine-knot structures —to serve as molecular dimerizers — may be particularly relevant to the 'tail' location of D3 in Beat.

Bork [15] has described a class of carboxy-terminal (CT) modules in molecules as varied as connective-tissue growth factor, Von Willebrand factor, several mucins and *Drosophila* slit protein, that show a clear resemblance to cysteine-knot folds.